Abstract
Liquid-liquid phase separation (LLPS) is a fundamental physicochemical process forming cellular condensates with a major role in the regulation of biomolecular processes such as DNA repair, cellular organization and regulation. Current efforts to engineer polypeptide-mediated LLPS rely on direct fusion (via protein linkers) of phase-separating peptides to proteins to be sequestered. For endogenous proteins, such approaches require genome editing that is difficult, time-consuming and can interfere with protein function/expression and genome architecture. This PhD project aims to address these limitations by introducing peptide arms (PeptArms), AI-designed binders fused to AI-generated phase-separating peptides, and showcasing their applications, for the versatile recruitment of both endogenous and exogenous proteins into condensates. This work leverages the lab expertise in AI-based peptide design, LLPS prediction, and synthetic biology, introducing a pioneering example of how advances in generative AI could address specific challenges in biomedicine, agriculture and biotechnology (in line with PotGenAI/SCAI call).

Project context and objectives
Liquid-liquid phase separation (LLPS) is a fundamental physicochemical process by which biomolecules such as proteins and RNAs demix from their surroundings to form dynamic, membrane-less organelles (*1*). These biomolecular condensates play important roles in various biological processes, including DNA repair, RNA metabolism, and photosynthesis (*2*, *3*). Condensation of a protein can occur in one or multiple steps: (i) self-LLPS, where a protein with phase-separating regions independently undergoes LLPS, or (ii) through the recruitment of non-phase-separating proteins via protein-protein interactions into condensates created by phase-separating partners. As an example, EPYC1 (Essential Pyrenoid Component 1), involved in the carbon concentrating mechanism of Chlamydomonas, phase-separates and recruits photosynthetic enzymes to form condensates known as pyrenoids (*4*). There is significant interest in engineering membrane-less organelles to control cellular processes for medical and environmental biotechnology applications (*5*). Traditional methods of fusing phase-separating peptides to proteins are challenging, time-consuming, and may disrupt protein expression and function due to the need for overexpression, gene knockout, or genome modification (*6*). Inspired by natural examples such as EPYC1 and leveraging advances in AI-based polypeptide design, this project introduces peptide arms (PeptArms). Peptide arms enable protein capture and condensation via AI-designed binders and (fused to) phase-separating peptides, respectively. This approach facilitates modular and versatile sequestration of both endogenous and exogenous proteins without requiring genome modifications. The proposed project here establishes the foundation for programmable bottom-up engineering of cellular assemblies for future research in medical, industrial and environmental applications. The PeptArms framework leverages the computational and experimental know-how of the lab to address three main objectives: **(1) Development and demonstration of AI-based PeptArms framework, (2) Reengineering of EPYC1 in Chlamydomonas using peptide arms, (3) Adapting the framework for endogenous and tag-free protein purification.**

Methodology (work packages, WPs, see figure below)
WP 1: Development of AI-based PeptArms framework (Months 1–12). We will establish AI-driven peptide arms (PeptArms) for targeted protein recruitment into biomolecular condensates. Computational design tools, including deep-learning-based binder modeling and LLPS prediction algorithms, will be used to optimize fusion constructs. Experimental validation in E. coli (on GFP and mCherry) will confirm condensate formation via fluorescence microscopy. **De novo binder design tools:** Four state-of-the-art models will be used to explore their potential in de novo binder design. BindCraft (*7*) and AlphaProteo (*8*) are based on AlphaFold2 weights, PepPrCLIP (*9*) leverages contrastive language-image pretraining to generate naturalistic peptides, and RFdiffusion (*10*) uses a diffusion model to design binders. **Deliverables:** AI-designed binders, validated peptide constructs, and fluorescence-based confirmation of condensates.

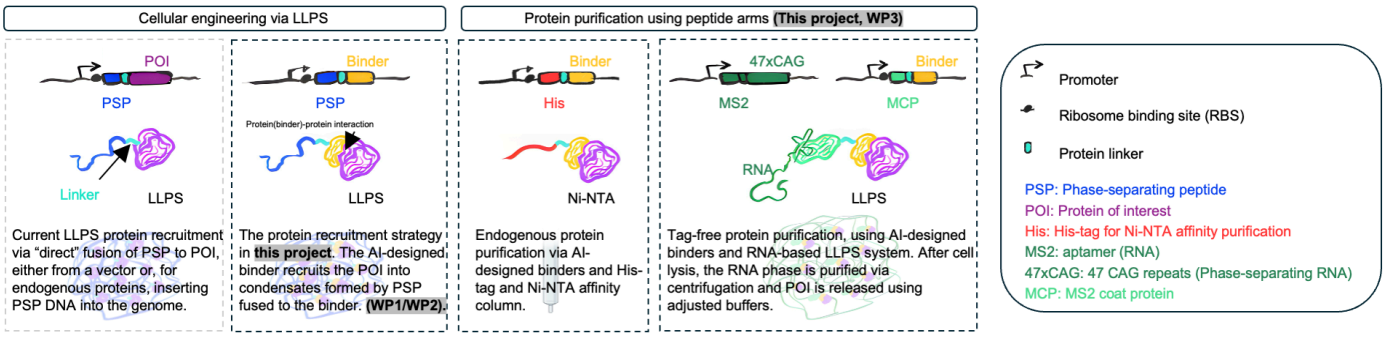
WP 2: Reengineering Synthetic Pyrenoid via PeptArms (Months 10–24). Using computational tools explained in WP1, synthetic EPYC1 analogs will be designed to include both LLPS and RuBisCo-recruitment modules. These analogs will be introduced into EPYC1-knocked out Chlamydomonas via a mutagenesis cassette, as described by Mackinder et al. (*4*). Engineering of the chloroplast genome of Chlamydomonas will be performed in collaboration with Lemaire and Crozet labs within the unit. Engineered strains will be assessed under varying CO_2 conditions for (i) the condensate morphology and packing density of RuBisCo and (ii) for metabolic efficiency and photosynthetic

performance. **Deliverables:** EPYC1 deficient Chlamydomonas strain, Synthetic EPYC1 designs, characterization of synthetic EPYC1 equivalents in condensate formation and photosynthesis efficiency.

WP 3: Adapting PeptArms framework for endogenous/tag-free protein purification (Months 22–36). Endogenous protein purification: For a target endogenous protein, binders will be designed using methods in WP1 and will be fused to a His-tag through a linker. The bound endogenous protein is purified using His-tag Ni-NTA affinity purification. Buffers will be adjusted for optimal purification and release of the target protein. This strategy can be used on lysate prepared from any cell type (prokaryotic or eukaryotic) for which specific binders fused to His-tag are purified beforehand. Tag-free protein purification using RNA-based LLPS: LLPS can also occur with phase-separating RNA forming condensates that can be separated from cell lysate via centrifugation (*11*). This enables tag-free protein purification. An RNA is transcribed that contains 47 CAG repeats enabling RNA LLPS fused to the MS2 aptamer (The co-director lab developed this technology in E. coli (*11*)). From a separate vector, specific binders targeting a protein of interest will be fused to MS2 coat protein (MCP). The binder binds to the target protein and is recruited into the RNA phase via the MCP-MS2 system (see project summary figure). The RNA phase will be separated by centrifugation after cell lysis. The bound protein will then be released from the RNA phase to yield purified, tag-free protein, validated using SDS-PAGE. **Deliverables:** Validated binders, optimized purification protocols, and functional protein recovery results.

Risk Mitigation: Potential binding inefficiencies will be addressed through iterative AI refinement, linker optimization, and alternative binder selection. Functional challenges in synthetic pyrenoid formation and protein purification will be managed through computational and experimental adjustments.

Supervisors role: The director (**A. Pandi**) will supervise the PhD student in a daily manner on technical aspects in binder design, and experimental work with support from co-director (**A. Lindner**) on the supervision of the student (since this is the first official supervision of the director, but received the temporary exemption from EDCdV for his ATIP-Avenir funding) and on scientific and technical aspects of the project on biophysics of liquid-liquid phase separation and RNA-based system proposed in the project. The feasibility of this project is supported by the host team expertise and ongoing research, both director and co-director, in synthetic biology (Pandi *et al*. 2019 ACS Synthetic Biology *(12)*, Pandi *et al*. 2019 Nature Communications *(13)*, Pandi *et al*. 2022 Nature Communications *(14)*), de novo peptide/DNA design (Pandi *et al*. 2023 Nature Communications *(15)*, Fallahpour *et al*. Nature Communications 2025 *(16)*) and LLPS-based cellular engineering via both RNA and peptides (Guo *et al*. 2022 Cell *(11)*, MohammadHosseini *et al*. 2025 bioRxiv *(17)*).



References

1. E. W. Martin, A. S. Holehouse, *Emerg Top Life Sci*. 4, 307–329 (2020). **2**. A. A. Hyman, C. A. Weber, F. Jülicher, *Annu Rev Cell Dev Biol*. 30, 39–58 (2014). **3**. Y. Gao, X. Li, P. Li, Y. Lin, *Nat Chem Biol*. 18, 1307–1318 (2022). **4**. L. C. M. Mackinder *et al.*, *Proc Natl Acad Sci U S A*. 113, 5958–5963 (2016). **5**. Z.-G. Qian, S.-C. Huang, X.-X. Xia, *Nat Chem Biol*. 18, 1330–1340 (2022). **6**. M. V. Garabedian *et al.*, *Nat Chem Biol*. 17, 998–1007 (2021). **7**. M. Pacesa *et al.*, *bioRxiv* (2024), doi:10.1101/2024.09.30.615802. **8**. V. Zambaldi *et al.*, *arXiv (*2024), doi:10.48550/ARXIV.2409.08022. **9**. S. Bhat *et al.*, *Sci Adv*. 11, eadr8638 (2025). **10**. J. L. Watson *et al.*, *Nature*. 620, 1089–1100 (2023). **11**. H. Guo, … **A. B. Lindner**, *Cell*. 185, 3823–3837.e23 (2022). **12. A. Pandi** *et al.* 2019 *ACS Synth Biol*, 23;8(8):1952-7 (2019). **13. A. Pandi** *et al. Nat Commun* 28;10(1):3880 (2019). **14. A. Pandi** *et al.* Nat Commun 5;13(1):3876 (2022). **15. A. Pandi** *et al. Nat Commun* 8;14(1):7197 (2023). **16.** A. Fallahpour, …, **A. B. Lindner, A. Pandi**, *bioRxiv* (2024) doi: 10.1101/2024.09.13.612903 (Accepted at Nature Communication Feb. 2025). **17.** A. M. MohammadHosseini …, **A. B. Lindner, A. Pandi**, *bioRxiv* (2025) doi: 10.1101/2025.01.27.635039.

The director and c-director names are in **bold**.